



TITLE:

方程式系の近似解に対する誤差評価 (数値計算のアルゴリズムの研究)

AUTHOR(S):

山本, 哲朗

CITATION:

山本, 哲朗. 方程式系の近似解に対する誤差評価 (数値計算のアルゴリズムの研究). 数理解析研究所講究録 1981, 422: 1-11

ISSUE DATE:

1981-04

URL:

<http://hdl.handle.net/2433/102565>

RIGHT:

方程式系の近似解に対する誤差評価

愛媛大 理 山本 昭朗

$x^{(0)}$ を \mathbb{R}^n または \mathbb{C}^n のある領域 D で定義した方程式

$$(1) \quad f(x) = (f_1(x), \dots, f_n(x))^T = 0$$

の近似解とするとき, その精度を判定することは, 数値解析における主要な課題の一つであって, 多くの文献があるが, そのらは通常適当なノルムを用いてなしている。しかし, $x^{(0)}$ の成分の絶対値に大きな差異がある場合, 通常のノルム評価は好ましくない。本稿では, 筆者が最近発表した一連の結果 [11], [12], [13] をより一般化し, 代数的問題に対する浮動小数点誤差解析の結果を数値例と共に報告した。

§1. 解の存在定理

次の記号を用いる。 $x = (x_i), y = (y_i) \in \mathbb{R}^n$ 或 \mathbb{C}^n に対し

$$\nu[x] = (|x_i|), \quad \rho(x, y) = \nu[x - y].$$

また, $x, y \in \mathbb{R}^n$ かつ $x_i \geq y_i$ ($\forall i$) のとき $x \geq y$ または $y \leq x$

とかく。行列, テンソルについても同じ記号を用いる。さらに

2

$v \geq 0$ に対し

$$U(x^{(0)}, v) = \{x \in \mathbb{R}^n (\text{or } \mathbb{C}^n) \mid \rho(x, x^{(0)}) \leq v\}$$

とおく. 以下において次のことを仮定する.

(i) f のヤコビ行列 $J(x) = (\frac{\partial f_i}{\partial x_j})$ は D 上存在する.

(ii) n 次行列 $L = (l_{ij})$ を

$$K = v[I_n - LJ(x^{(0)})] = (k_{ij})$$

のスペクトル半径 $\rho(K) < 1$ とするものが存在する.

(iii) 3次の対称テンソル $H = (h_{ijk}(D))$ を

$$(2) \quad \rho(LJ(x), LJ(y)) \leq H \rho(x, y), \quad x, y \in D$$

となるものが存在する.

このとき, 次の定理が成り立つ.

定理 1. $\|\cdot\|$ を任意のノルムとし, $\|x\| = \|y\| = 1$ なるすべての

x, y につき

$$(3) \quad Kx \leq \kappa, \quad Hxy \leq h.$$

かつ $\|\kappa\| < 1$ とする. さらに

$$\varepsilon = v[Lf(x^{(0)})], \quad t = (1 - \|\kappa\|)^2 - 2\|h\|\|\varepsilon\| \geq 0$$

とすると

$$a = \frac{2\|\varepsilon\|}{1 - \|\kappa\| + \sqrt{t}}, \quad \beta = \varepsilon + a\kappa + \frac{1}{2}a^2h = (\beta_1, \dots, \beta_n)^t$$

とおく. 以下

(i) $U(x^{(0)}, \beta)$ 内には (1) の解 x^* が存在する.

(ii) $\{\beta^{(k)}\} \in$

$$\beta^{(0)} = \beta, \quad \beta^{(k+1)} = \varepsilon + \kappa \beta^{(k)} + \frac{1}{2} H \beta^{(k)^2}, \quad k \geq 0$$

により定義可なり、 $\beta^{(0)} \geq \beta^{(1)} \geq \dots \rightarrow \beta^* \geq 0$ かつ $x^* \in U(x^{(0)}, \beta^*)$.

(iii) x^* は $U(x^{(0)}, \beta^*)$ 内で一意.

(iv) 特にな $\varepsilon > 0$ ならば x^* は単純である.

(証明). Schröder, Collatz による majorant principle による.

詳細は山本 [14].

系. 定理1の仮定の下で, $x^{(1)} = x^{(0)} - Lf(x^{(0)})$ とおけば

$$x^* \in U(x^{(1)}, \beta^* - \varepsilon).$$

注意1. ノルムとして l_1 ノルム $\|\cdot\|_1$ と最大ノルム $\|\cdot\|_\infty$ をとれば

$\kappa = (\kappa_i)$, $h = (h_i)$ はそれぞれ

$$\|\cdot\|_1: \quad \kappa_i = \max_j \kappa_{ij}, \quad h_i = \max_{j,k} h_{ijk}(D)$$

$$\|\cdot\|_\infty: \quad \kappa_i = \sum_{j=1}^m \kappa_{ij}, \quad h_i = \sum_{j,k=1}^m h_{ijk}(D)$$

により与えられる.

注意2. 定理1は最大ノルムにつき, Newton-Kantorovich の定理を改良している. 実際, $f: \kappa \cdot D$ において C^2 級とし

$$\left| \frac{\partial^2 f_i}{\partial x_j \partial x_k} \right| \leq m_{ijk} \quad (\forall i, j, k), \quad M = (m_{ijk})$$

とすれば, 注意1により

$$\|\kappa\|_\infty = \|\kappa\|_\infty, \quad \|h\|_\infty = \|H\|_\infty \leq \|L\|_\infty \|M\|_\infty.$$

故に $\tilde{\varepsilon} = (1 - \|\kappa\|_\infty)^2 - 2\|L\|_\infty \|M\|_\infty \geq 0$ ならば

$$t_\infty = (1 - \|\kappa\|_\infty)^2 - 2\|h\|_\infty \|\varepsilon\|_\infty \geq 0$$

で

$$\|x^{(0)} - x^*\|_\infty \leq \|\beta\|_\infty \leq \alpha_\infty = \frac{2\|\varepsilon\|_\infty}{1 - \|K\|_\infty + \sqrt{\kappa_\infty}} \leq \frac{2\|\varepsilon\|_\infty}{1 - \|K\|_\infty + \sqrt{\varepsilon}}$$

ここで特に $L = J(x^{(0)})^{-1}$ とおけば $K = 0$ で、上式最後の値は Newton-Kantorovich の評価と一致する。

さて、定理1の誤差限界ベクトル β^* は L に依存している。
では、どのような L に対し最良の β^* がえられるであろうか。
当然予想される結果は

$$\beta_{J(x^{(0)})^{-1}}^* \leq \beta_L^* \quad (V_L)$$

である。実際、この予想は、ある条件下で正しい。これを示すために、 $K, \varepsilon, \kappa, \beta^*$ 等を $K_L, \varepsilon_L, \kappa_L, \beta_L^*$ 等と置き、特に $L = J(x^{(0)})^{-1}$ に対するものを $\hat{K}, \hat{\varepsilon}, \hat{\kappa}, \hat{\beta}^*$ 等によりあらわす。

定理2. H_L は次の意味で最良とする。すなわち $H_L' \geq 0$ を

$$\rho(LJ(x), LJ(y)) \leq H_L' \rho(x, y), \quad x, y \in D$$

をみたす3次の対称テンソルとすると、 $H_L \leq H_L'$ 。もし

$\|\cdot\|$ が単調で $\|K_L\| < 1$ かつ $\kappa_L = (1 - \|K_L\|)^2 - 2\|K_L\| \cdot \|\varepsilon_L\| > 0$ ならば

$$\hat{\kappa} = 1 - 2\|\hat{K}\| \cdot \|\hat{\varepsilon}\| > 0 \quad \text{かつ} \quad \hat{\beta}^* \leq \beta_L^*.$$

(証明) 山本 [14].

§2. 応用

2.1. 連立一次方程式

定理3. $x^{(0)} \in Ax=b$ の近似解, $L \in A^{-1}$ の近似とし

$$\|K\|_\infty < 1 \quad (\text{但し } K = \nu[I_n - LA] = (K_{ij}))$$

と収束する。

$$\varepsilon = \nu[L(Ax^{(0)} - b)], \quad \kappa_i = \sum_{j=1}^n \kappa_{ij}, \quad \kappa = (\kappa_1, \dots, \kappa_n)^T$$

$$\beta^{(0)} = \varepsilon + \frac{\|\varepsilon\|_\infty}{1 - \|\kappa\|_\infty} \kappa, \quad \beta^{(k+1)} = \varepsilon + \kappa \beta^{(k)}, \quad k \geq 0$$

とわかる。 $\beta^{(0)} \geq \beta^{(1)} \geq \dots \rightarrow \beta^* = (I_n - \kappa)^{-1} \varepsilon$ かつ

$$|x_i^* - x_i^{(0)}| \leq \beta_i^* \leq \beta_i^{(k)}, \quad i=1, 2, \dots, n \quad (k \geq 0).$$

2.2. 固有値問題

実行列 A の近似実固有値 $\lambda^{(0)}$ と近似実固有ベクトル $x^{(0)}$ が与えられる。 $z^{(0)} = (x^{(0)}, \dots, x_n^{(0)}, \lambda^{(0)})^T$ と

$$f(z) = \begin{pmatrix} f_1(z) \\ \vdots \\ f_m(z) \end{pmatrix} = \begin{pmatrix} Ax - \lambda x \\ \frac{1}{2}(1 - \|x\|_2^2) \end{pmatrix} = 0$$

の近似解とみなせる。この場合、 $m = n+1$ として

$$J(z^{(0)}) = \begin{pmatrix} A - \lambda^{(0)} I_n & -x^{(0)} \\ -x^{(0)T} & 0 \end{pmatrix}$$

$$H = \nu \left[\begin{array}{c|c|c} l_{1m} & 0 \dots 0 & l_{1n} \\ l_{2m} & 0 \dots 0 & l_{2n} \\ \vdots & \vdots & \vdots \\ l_{mm} & 0 \dots 0 & l_{mn} \end{array} \middle| \begin{array}{c|c|c} 0 & l_{1m} & 0 \dots 0 \\ 0 & l_{2m} & 0 \dots 0 \\ \vdots & \vdots & \vdots \\ 0 & l_{mm} & 0 \dots 0 \end{array} \middle| \begin{array}{c|c|c} 0 & \dots & 0 \\ 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & 0 \end{array} \right]$$

に注意可。次の結果を得る。

定理4. $\|\cdot\|$ を任意のノルムとし、 L は $J(z^{(0)})^{-1}$ の近似と可る。

$$\kappa = \nu[I_{m+1} - L J(z^{(0)})], \quad \varepsilon = \nu[L f(z^{(0)})]$$

とし、 κ, h を定理1のよう定義する。 $\|\kappa\| < 1$ かつ $t = (1 - \|\kappa\|)^2$

$-2\|\kappa\| \|\varepsilon\| \geq 0$ と仮定して

$$a = \frac{2\|\varepsilon\|}{1 - \|\kappa\| + \sqrt{t}}, \quad \beta = \varepsilon + a\kappa + \frac{1}{2}a^2 h = (\beta_1, \dots, \beta_{m+1})^T,$$

$$\beta^{(0)} = \beta, \quad \beta^{(k+1)} = \varepsilon + \kappa \beta^{(k)} + \frac{1}{2} H \beta^{(k)2}, \quad k \geq 0$$

とおり、実固有値 λ^* と実固有ベクトル x^* ($\|x^*\|_2 = 1$) が存在して

$$\begin{pmatrix} x^* \\ \lambda^* \end{pmatrix} \in U(x^{(0)}, \beta^*).$$

すなわち

$$|x_i^* - x_i^{(0)}| \leq \beta_i^* \leq \beta_i^{(k)}, \quad i=1, 2, \dots, n$$

$$|\lambda^* - \lambda^{(0)}| \leq \beta_{n+1}^* \leq \beta_{n+1}^{(k)}, \quad k \geq 0.$$

この結果は、すべての固有値、固有ベクトルを必要としないこと、 A の対称性を仮定しないこと等、従来のものより実用的と考えられる。

§3. 誤差解析

実計算では、 $\beta^{(k)}$ は正しく計算されないから、定理 3, 4 を直接応用するには幾分の配慮が必要となる。いま演算結果を N 進浮動小数の仮数部に切捨てる計算機を考え、Wilkinson [9], Paige [4] に従って、 $n \geq 2$, か

$$(4) \quad fl(\hat{a} \cdot \hat{b}) = \hat{a} \cdot \hat{b} (1 + \xi), \quad |\xi| < N^{1-n}$$

$$(5) \quad fl(\sqrt{\hat{a}}) = \sqrt{\hat{a}} (1 + \eta), \quad |\eta| < 1.01 N^{1-n}$$

$$(6) \quad 1.006(n+1)N^{1-n} < 0.01$$

を仮定する。ただし \hat{a}, \hat{b} は a, b の N 進浮動小数表現とし、 \circ は四則演算の一つをあらわすものとす。式 c, d, \dots についても、その演算結果を \hat{c}, \hat{d}, \dots とあらわす。

3.1. 定理 3 に対する誤差解析 (山本 [16])

簡単のため $\hat{A} = A$, $\hat{b} = b$, $\hat{x}^{(0)} = x^{(0)}$ とする。また L は A^{-1} に対する計算結果ととり, $\hat{L} = L$ とする。このとき, (4)と(6)を用いて定理3にあらわした諸量 $\hat{\kappa}$, $\hat{\varepsilon}$, $\hat{\alpha}$, ... を評価する。結果は次の通り。

定理5. $A = (a_{ij}) (= \hat{A})$, $\theta_n = n N^{1-A}$

$$\hat{A}_\infty = \max_i \text{pl}(\sum_{j=1}^n |a_{ij}|), \quad \hat{L}_\infty = \max_i \text{pl}(\sum_{j=1}^n |l_{ij}|),$$

$$\hat{\kappa}_\infty = \max_i \text{pl}(\sum_{j=1}^n \hat{\kappa}_{ij}) = \|\hat{\kappa}\|_\infty,$$

$$\hat{C}_\infty = 1.02 \hat{A}_\infty \|\hat{x}^{(0)}\|_\infty + 0.502 \|\widehat{A\hat{x}^{(0)}} - b\|_\infty$$

$$\hat{d}_\infty = 1.02 \hat{L}_\infty (\|\widehat{A\hat{x}^{(0)}} - b\|_\infty + \hat{C}_\infty)$$

$$\hat{e}_\infty = 1.03 (\hat{L}_\infty \hat{A}_\infty + \hat{\kappa}_\infty) + 0.502 \max_i \hat{\kappa}_{ii}$$

とおく。以下

$$\hat{e}_\infty \theta_n < 1 \quad \text{から} \quad \hat{\kappa}_\infty < 1 - \frac{1}{m} - \hat{e}_\infty \theta_n \quad (m > 0)$$

より

$$\hat{f}_\infty = 1.004(m-1) \left\{ \left(1 + \frac{1}{m} + m \hat{e}_\infty \theta_n\right) \hat{d}_\infty + (2n^{-1} + m \hat{e}_\infty) \|\hat{\varepsilon}\|_\infty \right\}$$

$$\Delta \hat{\beta}_i = 1.004 N^{1-A} \hat{\beta}_i + \left\{ \hat{d}_\infty (1 + m \hat{e}_\infty \theta_n) + m \|\hat{\varepsilon}\|_\infty \hat{e}_\infty + \hat{f}_\infty + \frac{1}{n} \hat{\alpha} \hat{\kappa}_\infty \right\} \theta_n$$

とおいて

$$|x_i^* - x_i^{(0)}| \leq \hat{\beta}_i + \Delta \hat{\beta}_i, \quad i=1, 2, \dots, n.$$

注意3. $\hat{L}_\infty \hat{A}_\infty$ は A の条件数を近似する量であり, この値が大きくなれば $\Delta \hat{\beta}_i$ は大きくなる, 定理3は有効に働く。また
(5)
方程式が悪条件の場合, 計算桁数 s をどのようにとるべきかを定理5は示唆している。

3.2. 定理4 に対し誤差解析. (山本[15]).

定理5 と同様, 次の結果が成り立つ.

定理6. $\beta_i < (1 + 1.03\theta_2 + 1.02\theta_{n+1})\hat{\beta}_i + \theta_{n+1}\delta_i$, $i=1, 2, \dots, n+1$

但し $\hat{\varepsilon} > 0$ を仮定し,

$$\delta_i = e_i + 1.004(ag_i + \alpha) + 2^{-1}(2a + \alpha\theta_{n+1})\alpha h_i,$$

$$a = 2d^{-1}\|\varepsilon\|_1, \quad d = 1 - \|\kappa\|_1 + \sqrt{\varepsilon},$$

$$\alpha = 2.02(d\hat{d})^{-1}\{(\|e\| + 1.4\|\varepsilon\|_1)d + (p + \omega + 1)\|\varepsilon\|_1\}$$

$$p = 1.02\text{fl}(g_1 + \dots + g_{n+1}) + 1.02\text{fl}(\hat{\kappa}_1 + \dots + \hat{\kappa}_{n+1}) + 1$$

$$g_i = \max_j g_{ij}$$

$$g_{ij} = \frac{1.004}{n+1}\{|l_{ij}| \cdot |\hat{h}_{ij}| \cdot (1 - \sigma_{j,n+1}) + \text{fl}(1 - \text{fl}(LJ)_{ij})\} \delta_{ij} + d_{ij}$$

$$(\hat{h}_{ij}) = \text{fl}(A - \lambda^{(0)}I_n)$$

$$d_{ij} = 1.02\text{fl}(|l_{i1}| + \dots + |l_{i,n+1}|)\|J^{(i)}\|_\infty \quad (J^{(i)}_1, \dots, J^{(i)}_{n+1} \text{ の } J^{(0)} \text{ の } i \text{ 行})$$

$$e = (e_1, \dots, e_{n+1})^t,$$

$$e_i = 1.02\{1.006\|\hat{f}\|_\infty + \frac{n}{n+1}\|c\|_\infty\}\text{fl}\left(\sum_{j=1}^{n+1}|l_{ij}|\right)$$

$$f = (f_1(x^{(0)}), \dots, f_n(x^{(0)}))^t,$$

$$c = (c_1, \dots, c_{n+1})^t, \quad c_i = \begin{cases} \text{fl}\left(\sum_{j=1}^{n+1}|\hat{h}_{ij}|\right)(1.02)\|x^{(0)}\|_\infty & (i \leq n) \\ \frac{1.004}{n}|\hat{f}_{n+1}| + 0.508\|x^{(0)}\|_2^2 & (i = n+1) \end{cases}$$

$$\omega = 1 + \frac{\tau}{\sqrt{\varepsilon}}, \quad \tau = \beta + 2.02\|\kappa\|_1\|e\| + 2$$

$$\beta = 2p + p^2\theta_{n+1} + 1$$

§4. 数値例

例1. Wilkinson の 5 点 1 次元条件付き方程式

$$0.876543x_1 + 0.617341x_2 + 0.589973x_3 = 0.863257$$

$$0.612314x_1 + 0.784461x_2 + 0.827742x_3 = 0.820647$$

$$0.317321x_1 + 0.446779x_2 + 0.476349x_3 = 0.450098$$

を FACOM 230-28 (単精度 16 進 6 桁) で Gauss 消去 (行交換) を用いて解くと

$$x^{(0)} = (0.6363233, -0.2946413E-1, 0.5486381)^T$$

を得る. A^{-1} の計算値 $L (= \hat{L})$ は $\hat{L}_\infty = 0.657 \dots E-5$ (山本 [13]) であった.

この場合, 倍精度 (16 進 14 桁) 計算により $\hat{K}_\infty, \hat{e}_\infty, \dots$ を求めて, 定理 5 を適用可小は

$$\hat{\beta} = \begin{pmatrix} 0.573591 \dots E-5 \\ 0.427810 \dots E-4 \\ 0.362315 \dots E-4 \end{pmatrix}, \quad \Delta\beta_i = 0.150 \dots E-9 < 0.151E-9.$$

$$\text{故に } \|x^* - x^{(0)}\| \leq \hat{\beta} + 0.151E-9 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \leq \begin{pmatrix} 0.574E-5 \\ 0.428E-4 \\ 0.363E-4 \end{pmatrix}.$$

例 2. 行列

$$A = \begin{pmatrix} 14 & 9 & 6 & 4 & 2 \\ -9 & -4 & -3 & -2 & -1 \\ -2 & -2 & 0 & -1 & -1 \\ 3 & 3 & 3 & 5 & 3 \\ -9 & -9 & -9 & -9 & -4 \end{pmatrix}$$

の近似固有値 $\lambda^{(0)} = 4.990$ と近似固有ベクトル $x^{(0)} = (0.707, -0.707, 0.0002, -0.0001, 0.0000)^T$ が与えられたとき, それらの誤差をいさべる.

定理 4 により $\hat{\kappa} = 0.799 \dots < 1$ かつ

$$\hat{\beta} = \begin{pmatrix} 0.1682048E-3 \\ 0.1686592E-3 \\ 0.2619984E-3 \\ 0.1619263E-3 \\ 0.6215292E-4 \\ 0.1005869E-1 \end{pmatrix}, \quad \hat{\beta}^{(1)} = \begin{pmatrix} 0.1084336E-3 \\ 0.1088856E-3 \\ 0.2015247E-3 \\ 0.1009613E-3 \\ 0.4388653E-5 \\ 0.1003479E-1 \end{pmatrix}$$

この場合, 定理 6 により

$$\beta_i \leq \hat{\beta}_i + 10^{-12}$$

を得る. (詳細は山本 [15]). 故に 定理 4 (6) は有効に働く.

References

- [1] Collatz, L.: Functional analysis and numerical mathematics, Academic Press 1966.
- [2] Kantorovich, L.V. and Akilov, G.P.: Functional analysis in normed spaces. Pergamon Press 1964.
- [3] Ortega, J.M. and Rheinboldt, W.C.: Iterative solution of nonlinear equations in several variables, Academic Press 1970.
- [4] Paige, C.C.: Error analysis of the symmetric Lanczos process for the eigenproblem, London Univ. Inst. of Computer Science, Tech. Note ICSI 209, 1969.
- [5] Rall, L.B.: Computational solution of nonlinear operator equations, John Wiley & Sons, Inc. 1969.
- [6] Schröder, J.: Nichtlineare Majoranten beim Verfahren der schrittweisen Näherung, Arch. Math. 7(1956), 471-484.
- [7] Schröder, J.: Über das Newtonsche Verfahren, Arch. Rat. Mech. Anal. 1(1957), 154-180.
- [8] Urabe, M.: A posteriori component-wise error estimation of approximate solutions to nonlinear equations, Lect. Notes in Computer Sci. 29,, Springer, 1975.
- [9] Wilkinson, J.H.: Rounding errors in algebraic processes, Prentice Hall 1963.
- [10] Wilkinson, J.H.: The algebraic eigenvalue problem, Oxford Univ Press 1965.
- [11] Yamamoto, T.: Componentwise error estimates for approximate solutions of nonlinear equations, JIP 2(1979), 121-126.
- [12] Yamamoto, T.: An existence theorem for solutions of nonlinear systems and its applications to algebraic equations, In 3rd USA-JAPAN computer conference proceedings, 300-304. AFIPS and IPSJ 1978.

- [13] Yamamoto,T.: Error bounds for computed eigenvalues and eigenvectors,
Numer.Math.34(1980), 189-199.
- [14] Yamamoto,T.: Error bounds for approximate solutions of systems of
equations, to appear.
- [15] Yamamoto,T.: Componentwise error estimates for approximate solutions of
systems of equations, to appear.
- [16] Yamamoto,T.: A posteriori componentwise error estimate for a computed
solution of a system of linear equations, to appear.